A simulation comparison of spatiotemporal and spatially-implicit size-structured models for northern shrimp and snow crab

Jie Cao, James T. Thorson, André Punt, Cody Szuwalski

CAPAM, La Jolla
Oct. 2, 2018

Photo: https://thenavigatormagazine.com/dfo-wins-controlling-agreement-legal-challenge/

Outline

- Spatially-implicit/Spatially-stratified/Spatiotemporal models
- Spatiotemporal modeling framework (snow crab/northern shrimp)
- Simulation experiments
  - recover spatial patterns/unbiased estimates of spatially-aggregated population quantities
  - implicitly accounts for movement processes
  - outperforms spatially-implicit models

Why spatiotemporal model ?

- Heterogeneous and complex spatial structure – population and fishery
- Spatially-implicit models – biased estimates of population quantities
- Spatially-stratified models
  - spatial strata/movement of individuals among strata

Correlations - process errors/fishery patterns

Spatial correlation (either based on adjacency or distance)

Spatiotemporal population model

- Combines theory and methods from population dynamics and geostatistics
- Assume population density varies continuously across space

$$x(s_i) \sim N\left(\frac{1}{|n_i|}\sum_{j \epsilon n_i} x(s_j), \sigma^2\right)$$

- Joint distribution for density at all locations
- Expand to account for size-structured population dynamics

Objectives

- Development – estimating population dynamics at a fine spatial scale
- Demonstration – two species (snow crab and northern shrimp)
- Comparison – spatiotemporal model vs. spatially-implicit model
- Evaluation – effect of sample size

Why size-structured models?

- Advantages:
  - Requires no ability to age animals (shrimps, crabs, lobsters)
  - Uses the data actually available
  - Vulnerability / maturity are often functions of size and not age

Process model

Abundance at size (n) for a given location s and time t

$$\boldsymbol{n}_{s,t+1} = g(\boldsymbol{n}_{s,t}) \circ e^{\boldsymbol{\varepsilon}_{s,t}}$$

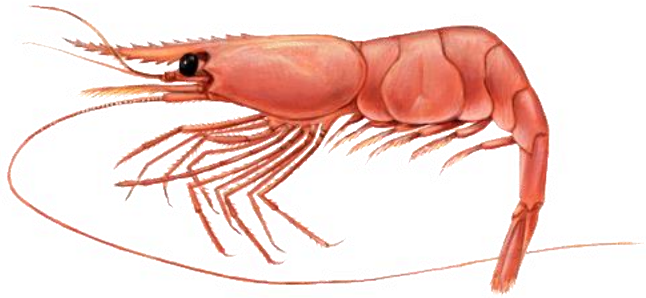
$$\boldsymbol{\Sigma}_t \sim \text{MVN}(0, \mathbf{R}_{spatial} \otimes \boldsymbol{\Theta}_L)$$

$\circ$	Hadamard product (entrywise product)
s	location
t	year
$\otimes$	Kronecker product

$\boldsymbol{n}_{s,t}$	vector of abundances for each of *l* size classes
$g()$	function representing population dynamic
$\boldsymbol{\varepsilon}_{s,t}$	vector of random effects (process error)
$\boldsymbol{\Theta}_L$	covariance among size classes (*l* by *l* matrix **L**)
$\mathbf{R}_{spatial}$	spatial covariance matrix (covariance between 2 locations follows a Matern function)

Population dynamic ($g()$)

- *Example 1: Gulf of Maine northern shrimp*



$$g(\mathbf{n}_{s,t}) = \mathbf{G}(\mathbf{n}_{s,t-1} \circ \exp(-\mathbf{m}_{s,t-1} - \mathbf{v}f_{s,t-1})) + \mathbf{r}_{s,t}$$

Initial condition:

$$g(\mathbf{n}_{s,1}) = \mathbf{r}_{s,t} \circ \exp(\boldsymbol{\varphi})$$

The predicted harvest per area:

$$\mathbf{c}_{s,t} = \frac{\mathbf{v}f_{s,t}}{\mathbf{v}f_{s,t} + \mathbf{m}_{s,t}} \circ \left(1 - \exp(-\mathbf{m}_{s,t} - \mathbf{v}f_{s,t})\right) \circ \mathbf{n}_{s,t}$$

$$\log(f_{s,t}) | \log(f_{s,t-1}) \sim N(\log(f_{s,t-1}), \sigma_f^2)$$

Population dynamic ($g()$)

- *Example 2: Eastern Bering sea snow crab*

$$g(\mathbf{n}_{s,t}^{\text{male}}) = \begin{cases} \mathbf{r}_{s,t}p^{\text{male}} + \mathbf{G}^{\text{male}}\left(\mathbf{n}_{s,t-1}^{\text{male}} \circ \exp\left(-\mathbf{m}_{s,t-1} - \mathbf{v}f_{s,t-1}^{\text{male}}\right)\right) \circ (1 - \mathbf{w}^{\text{male}}), & n = \mathbf{n}^{\lambda} \\ \mathbf{G}^{\text{male}}\left(\mathbf{n}_{s,t-1}^{\text{male}} \circ \exp\left(-\mathbf{m}_{s,t-1} - \mathbf{v}f_{s,t-1}^{\text{male}}\right)\right) \circ \mathbf{w}^{\text{male}} + \mathbf{n}_{s,t-1}^{h} \circ \exp\left(-\mathbf{m}_{s,t-1} - \mathbf{v}f_{s,t-1}^{\text{male}}\right), & n = \mathbf{n}^{\omega} \end{cases}$$

Initial condition: $g(\mathbf{n}_{s,1}^{\text{male}}) = \mathbf{r}_{s,1}p^{\text{male}} \circ \exp(\boldsymbol{\varphi}_{\text{male}})$

The predicted harvest per area: $\mathbf{c}_{s,t} = \left(1 - \exp\left(-\mathbf{v}f_{s,t}^{\text{male}}\right)\right) \circ \mathbf{n}_{s,t}^{\text{male}} \circ \exp\left(-0.5\mathbf{m}_{s,t}\right)$

$$\log(f_{s,t}) | \log(f_{s,t-1}) \sim N(\log(f_{s,t-1}), \sigma_f^2)$$

Model parameters and estimation

Fixed effects

$\boldsymbol{\Theta}_L$	process error covariance (among size classes)
κ	geostatistical range for correlations
μ_t	average offset of annual recruitment
φ	initial abundance of each size class
s	parameters of selectivity (logistic)
	Parameters of observation model

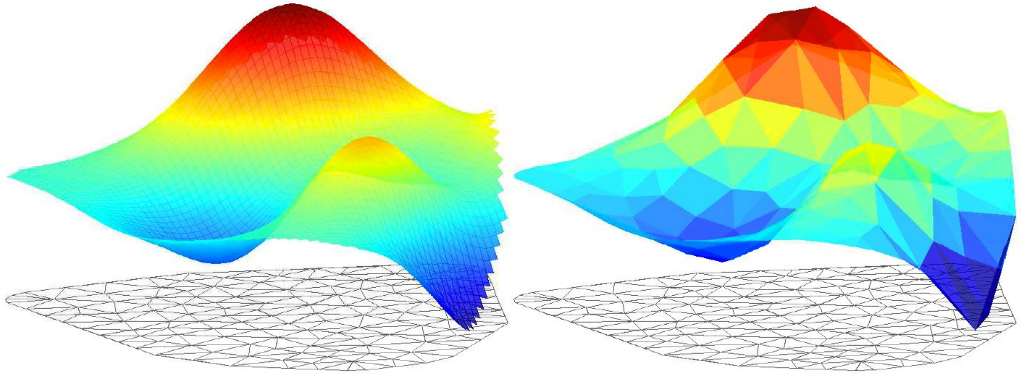
Random effects

r_t^u	spatial variation in recruitment
n_t	spatial variation in density for each size class and year
f	fishing mortality of location s over time

treat density as random, rather than process errors (ε_t)

Gaussian Markov random field (GMRF)

- Continuous spatial process -> discretely indexed GMRF
- Matérn covariance function
- Mesh/knot
- SPDE – MVN
- Piecewise constant
- Catch – lognormal
- Survey – lognormal/Poisson-link



Thorson, J.T., Shelton, A.O., Ward, E.J. and Skaug, H.J., 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. *ICES Journal of Marine Science*, *72*(5), pp.1297-1310.

Input data

survey data

Size_class	Year	Catch_N	AreaSwept_km2	Vessel	Lat	Lon
1	1	553	3.1	0	60	-174
1	1	629	3.1	0	63.5	-172
1	1	575	3.1	0	58	-170
1	1	618	3.1	0	61.5	-178
1	1	625	3.1	0	64.5	-170
1	1	634	3.1	0	61	-172

- used to create mesh/knots

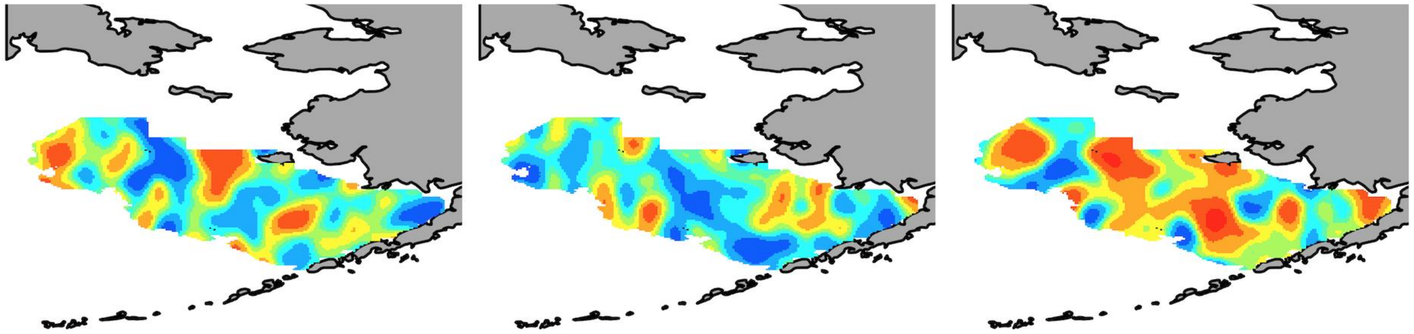
commercial catch data

X	lat	lon	year	X.1	X.2	X.3	X.4	X.5
1	55.6	-169	1	0.802	2.82	2.32	3.18	7.18
2	56.3	-170	1	0.657	1.83	1.54	2.15	4.94
3	56.3	-170	1	0.662	1.82	1.54	2.16	4.96
4	56.2	-171	1	0.64	1.78	1.5	2.1	4.81
5	56.2	-170	1	0.645	1.8	1.51	2.12	4.85
6	56.2	-170	1	0.646	1.82	1.52	2.13	4.88

- fine scale
- aggregated to knot-level

Model outputs

- Predicted population density map
- Estimated fishing mortality map
- Predicted catch map
- Estimated covariance of process error



Simulation – operating model

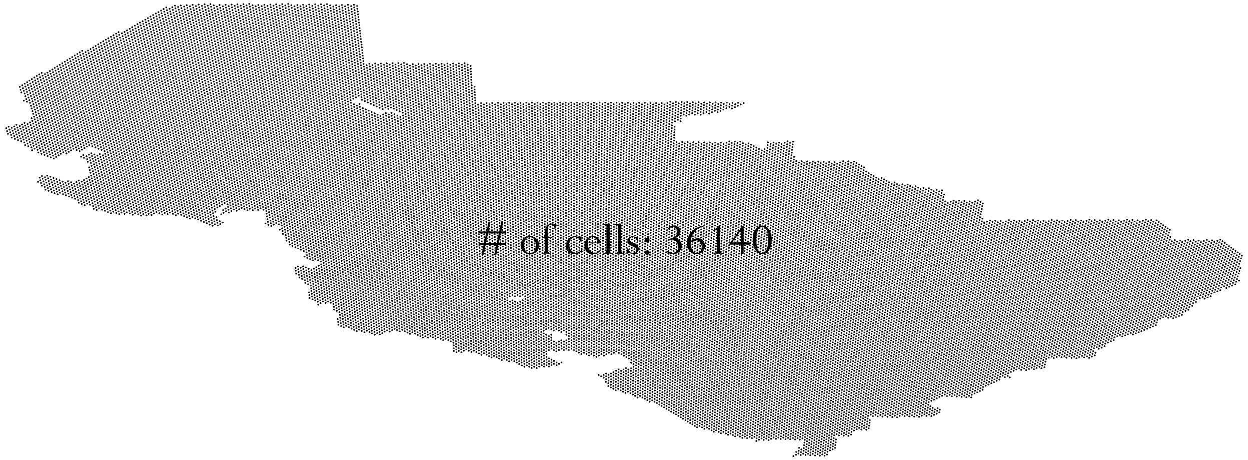
- Dynamics occur at fine scale
- Population dynamics (non-spatial) formulated identically to EM
- Cell-specific parameters (spatially correlated)
- Annual time step
- Movement $\mathbf{N}_{t+1} = g(\mathbf{M}\mathbf{N}_t) \circ e^{\mathbf{\Sigma}_t}$

$$\frac{\partial}{\partial t}\mathbf{n} = \mathbf{N}\mathbf{n}$$

***N** is the matrix of instantaneous movement rates*

$$\mathbf{M} \approx \left(\mathbf{I} + \frac{\mathbf{N}\Delta t}{n_{tdiv}}\right)^{n_{tdiv}}$$

***M** is annual movement rates*



Simulation experiments

1. Explore how the spatiotemporal model performs when individual movement processes are modeled explicitly
2. Compare spatially-implicit and spatiotemporal models
3. Evaluate the impact of changing sample size

Simulation experiments

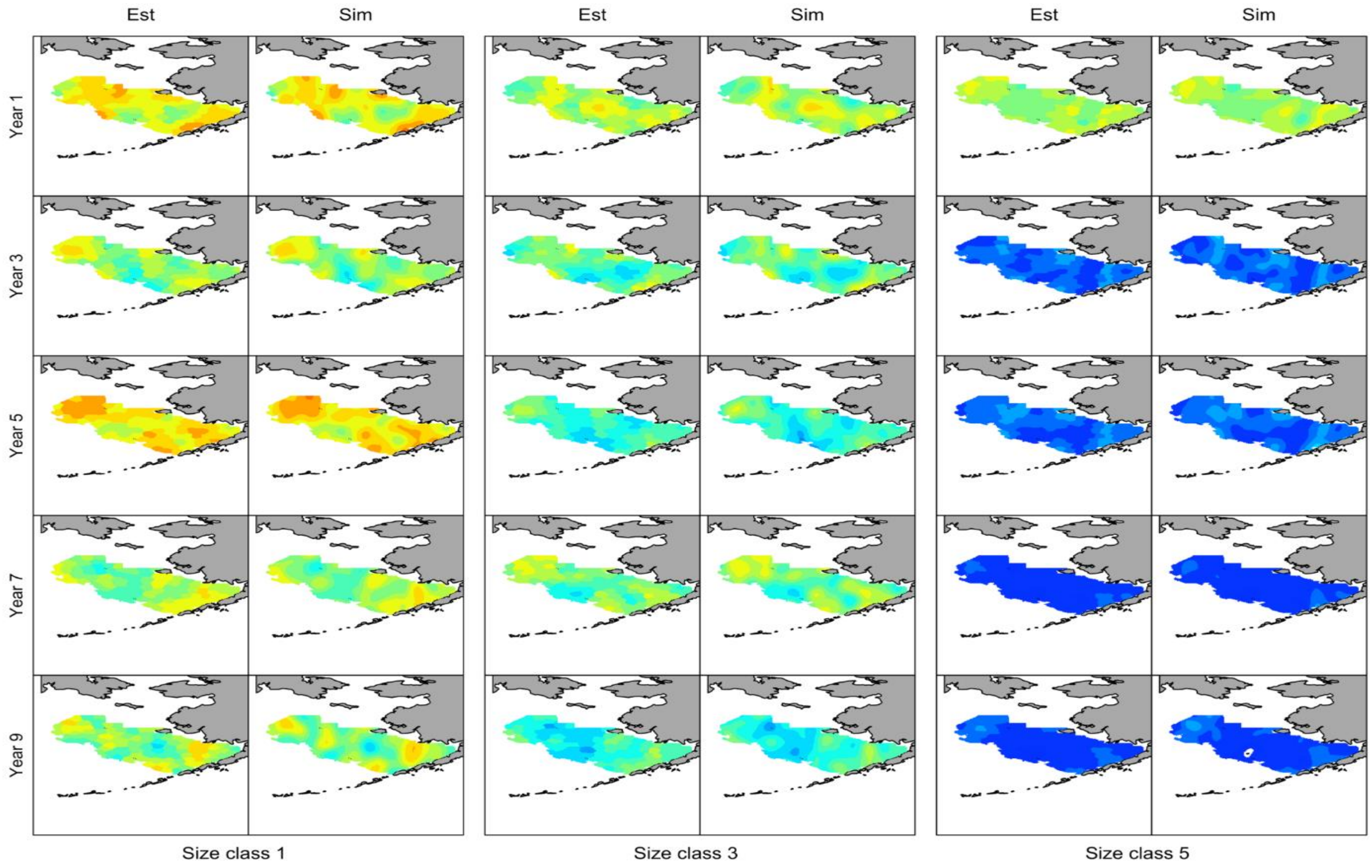
— Experiment 1: Exploring movement

1. No measurement error and no movement in the OM
2. Same as scenario 1, except there is movement
3. Both measurement error and movement in the OM

- 200 sites (grid cells) in the OM were randomly sampled each year
- For each site, total abundance by size class and the total area of the sampled site were recorded
- Fishery catch-at-size was calculated at each of the 36,140 grids and then aggregated to the knot level as data for the EM
- For the scenarios with measurement error, we generated 100 replicated data sets with sampling errors, i.e., grid-based survey abundance and fishery catch data were assumed to be lognormally distributed

Simulation experiments

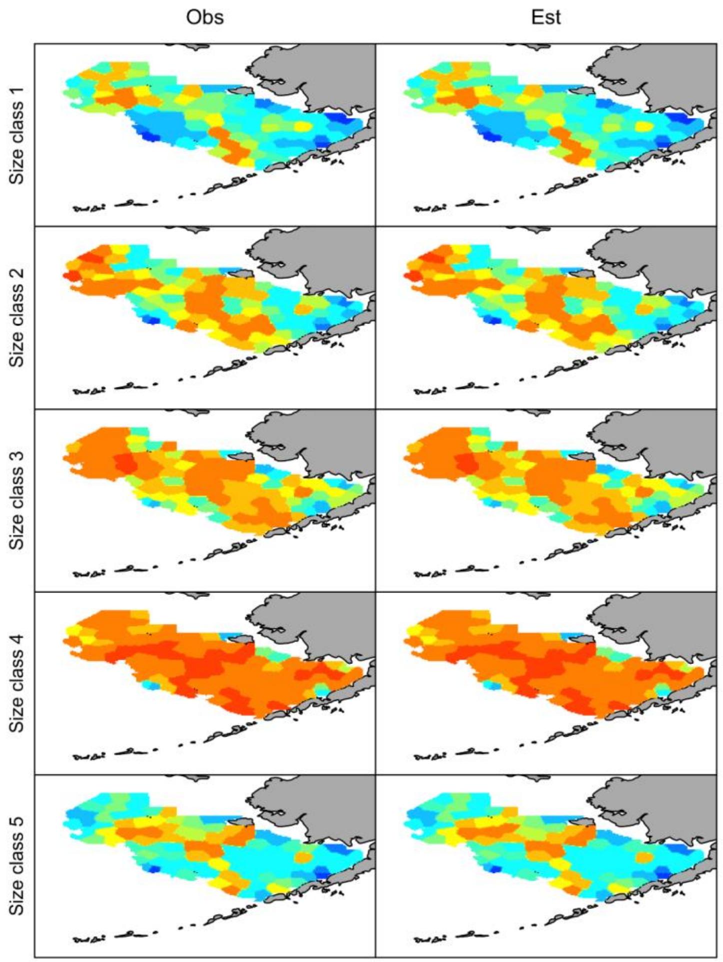
— Experiment 1: Exploring movement



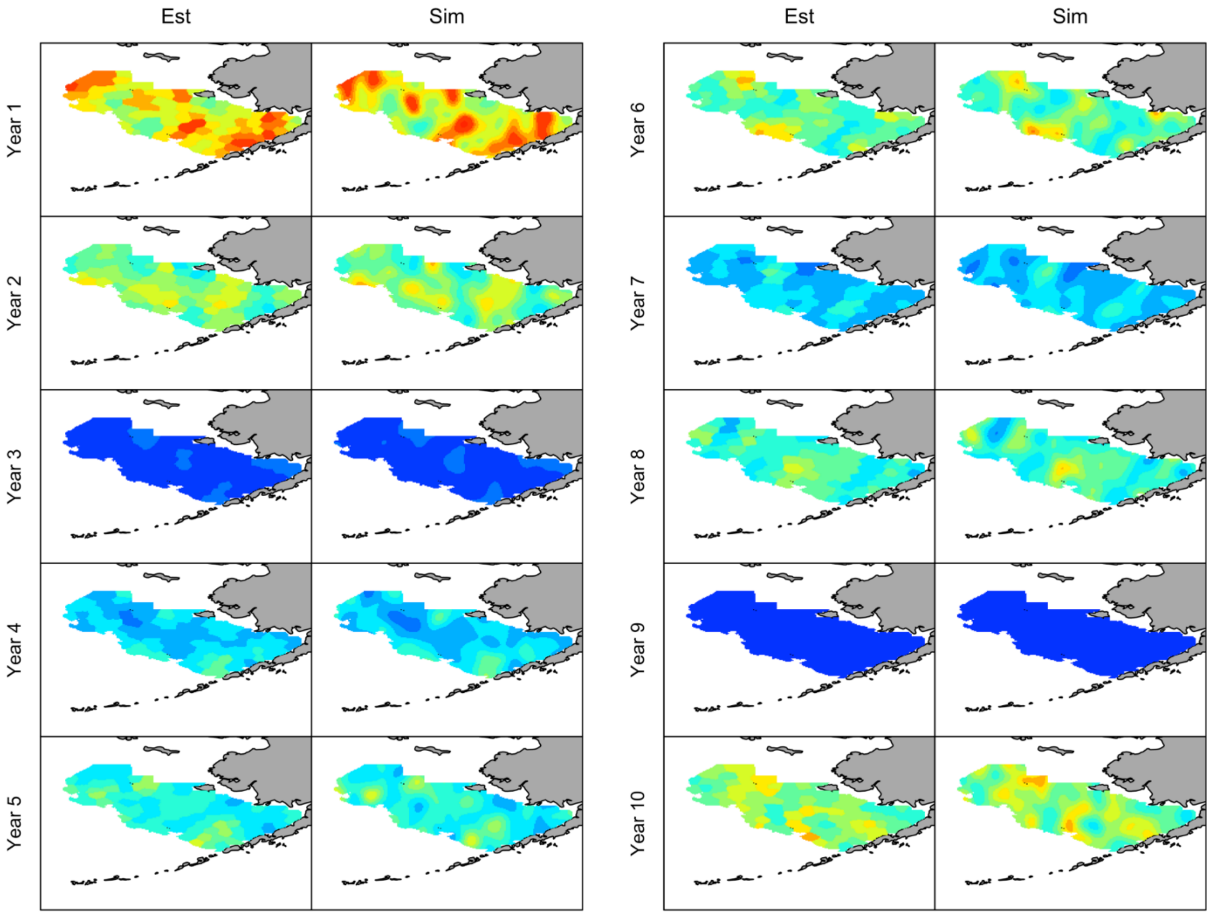
Simulation experiments

— Experiment 1: Exploring movement

Catch



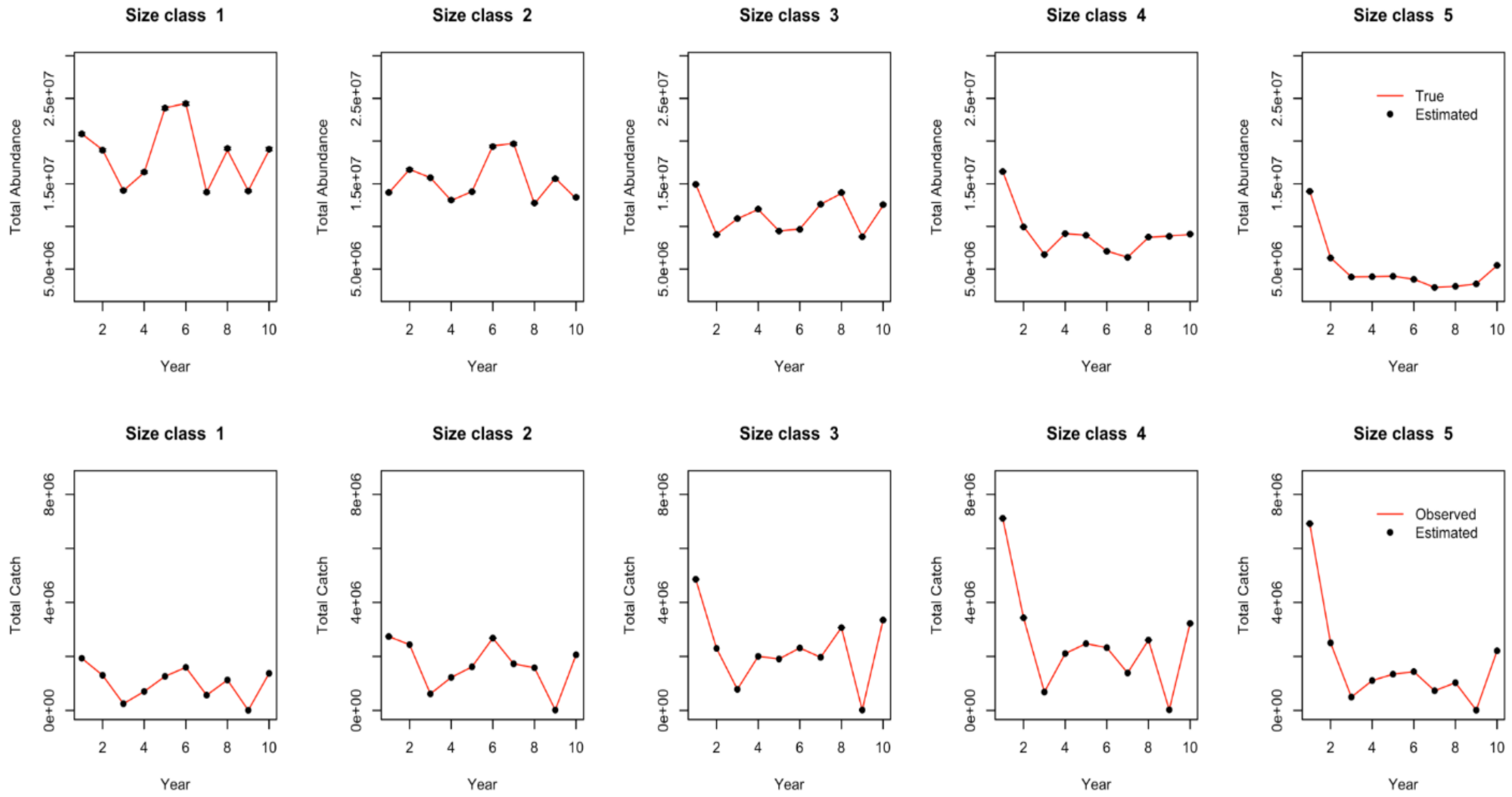
Fishing mortality



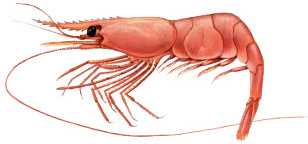
Simulation experiments

— Experiment 1: Exploring movement

Spatially-aggregated total abundance (a) and total removals (b) by size class over time

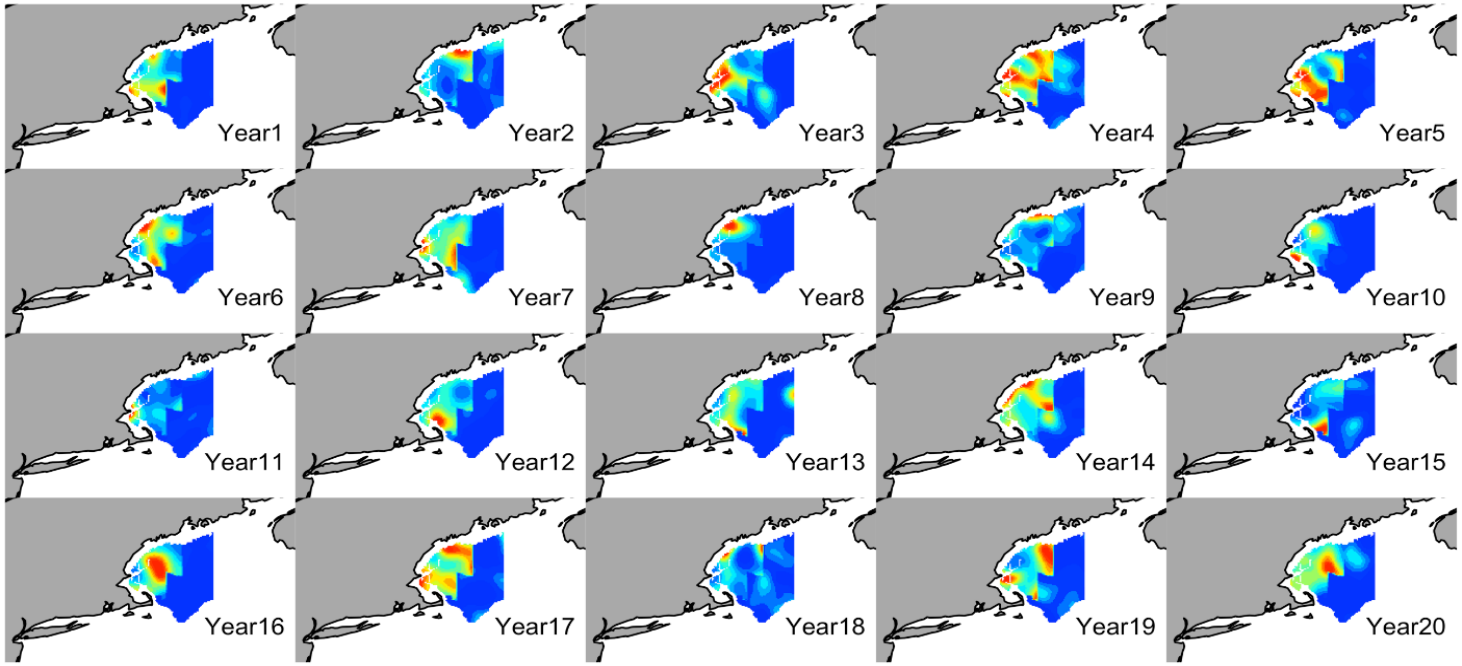


Simulation experiments



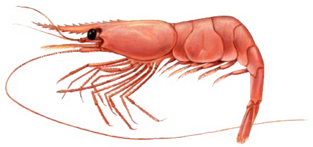
– Experiment 2: comparison of spatiotemporal and spatially-implicit models

Spatially-implicit model – size structured assessment model for Pandalus (Cao et al. 2017)



Simulated fishing mortality for northern shrimp (inshore area has persistent higher fishing mortality over time than offshore area)

Simulation experiments



– Experiment 2: comparison of spatiotemporal and spatially-implicit models

- the data used in both estimation models are the same at the grid spatial scale
- 50 knots for the spatiotemporal model
- a metric that is directly comparable
- abundance-at-size, fishing mortality at size and spawning stock biomass aggregated over the spatial domain

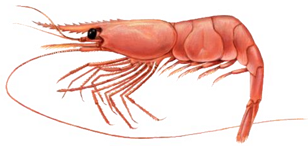
- **population-level fishing mortality**
- **aggregate selectivity-at-length**

$$c_{l,t} = \left(1 - \exp(-s_{l,t} f_t)\right) n_{l,t} \exp(-m_t)$$

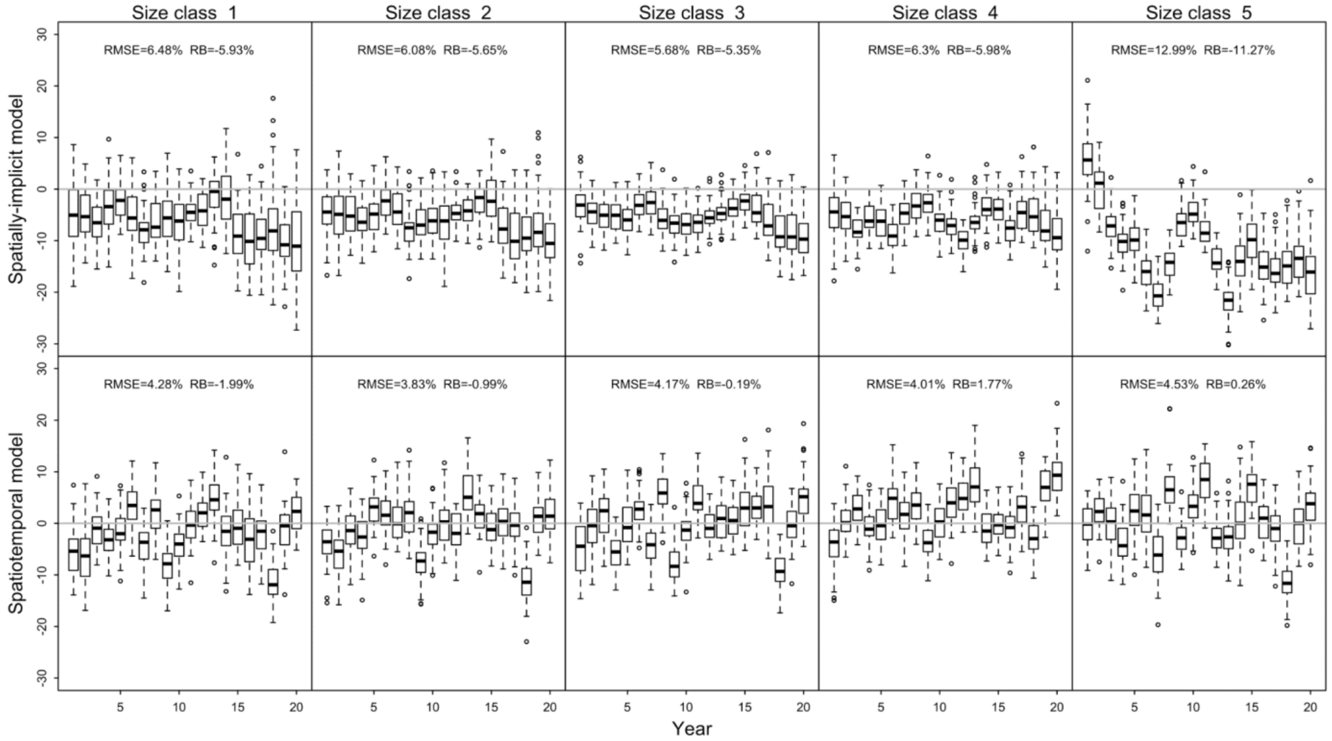
$$\mathrm{RMSE}_l = \sqrt{\frac{\sum_t \left(n_{l,t}^{est} - n_{l,t}^{true} \Big/ n_{l,t}^{true}\right)^2}{\tau}} \times 100\%$$

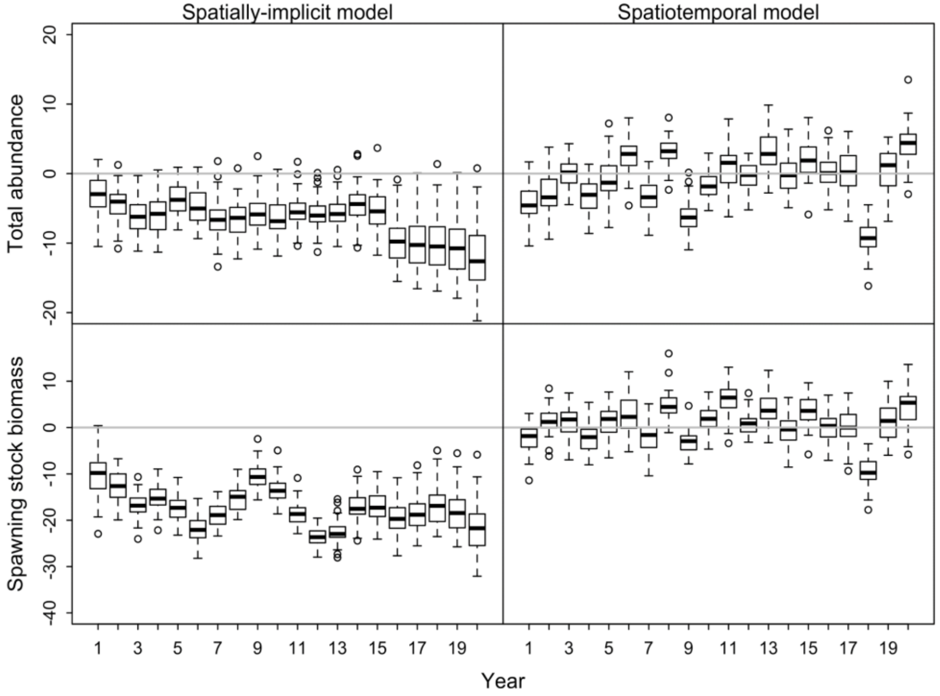
$$\mathrm{RB}_l = \frac{\sum_t \left(n_{l,t}^{est} - n_{l,t}^{true} \Big/ n_{l,t}^{true}\right)}{\tau} \times 100\%$$

Simulation experiments



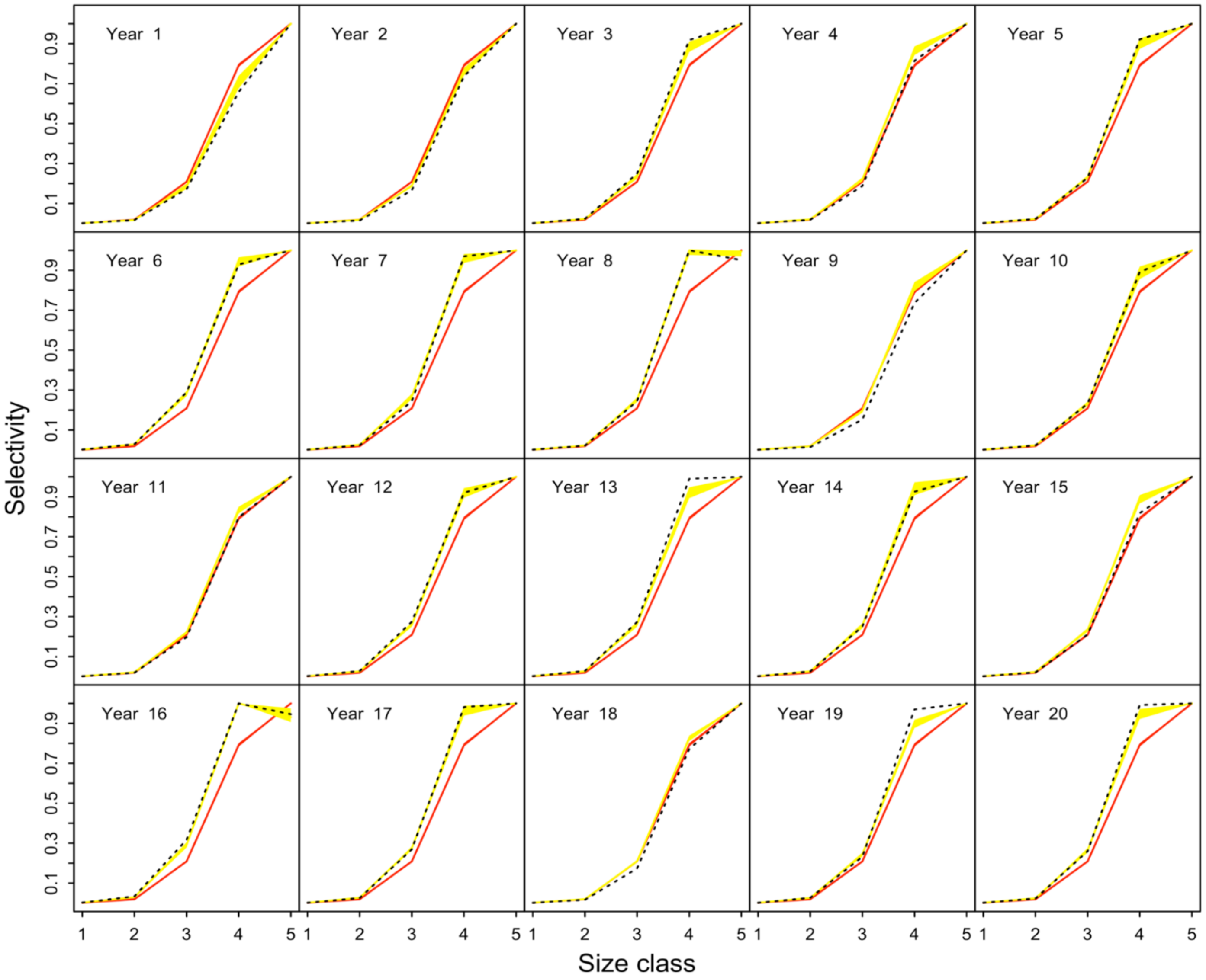
– Experiment 2: comparison of spatiotemporal and spatially-implicit models





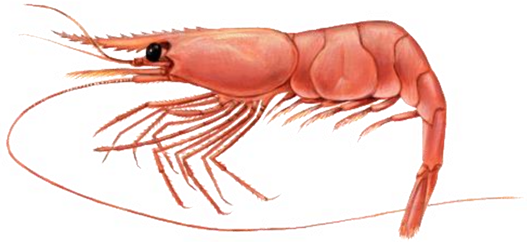
Simulation experiments

– Experiment 2: comparison of spatiotemporal and spatially-implicit models



Simulation experiments

— Experiment 3: Effect of sample size



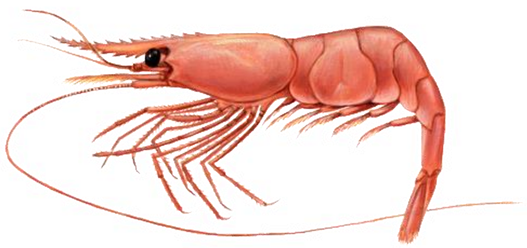
data poor	50 locations
moderate level	100 locations
data rich	200 locations

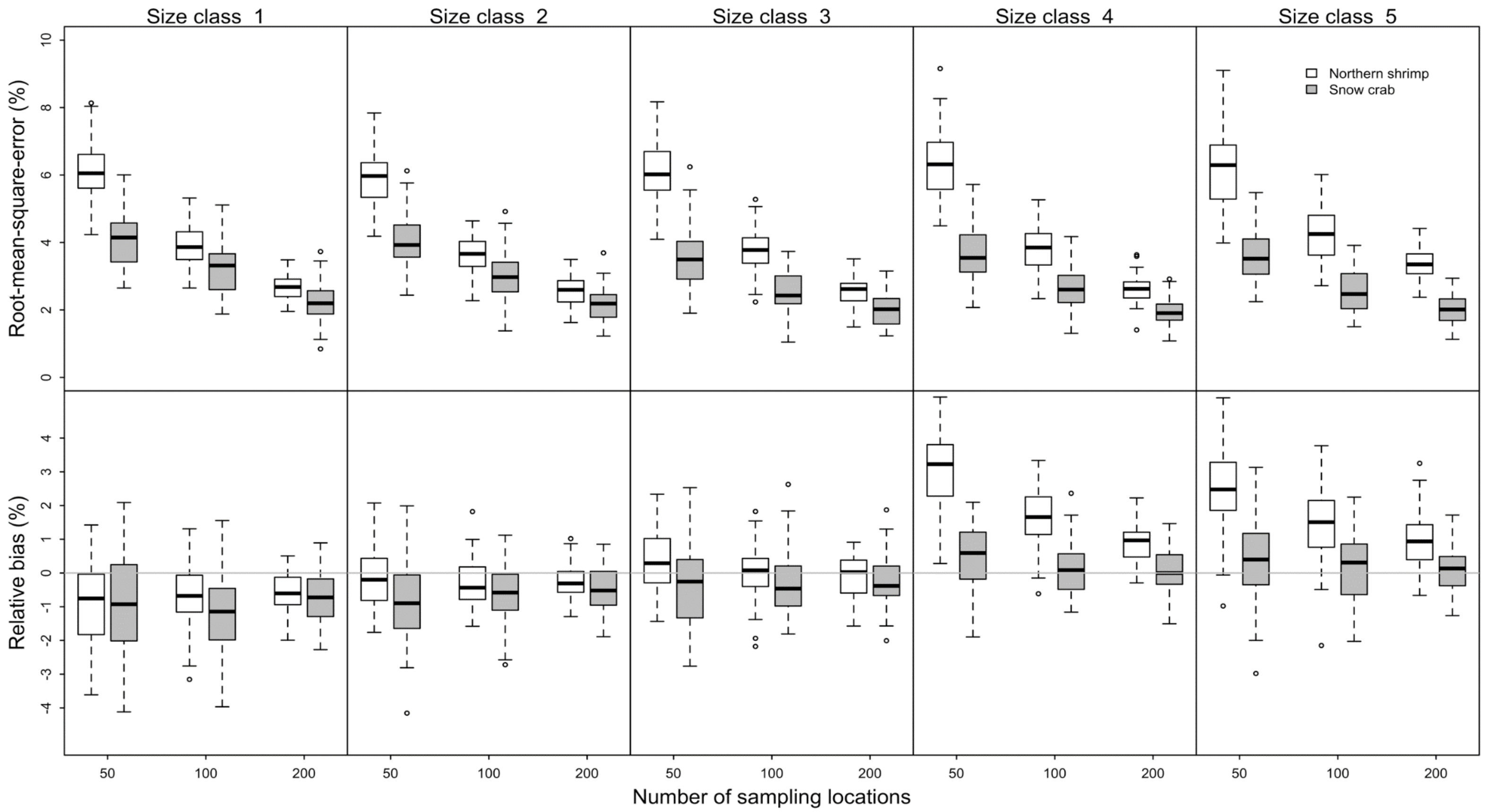
$$\mathrm{RMSE}_l = \sqrt{\frac{\sum_t \left({n_{l,t}^{est} - n_{l,t}^{true}} \Big/ {n_{l,t}^{true}} \right)^2}{\tau}} \times 100\%$$

$$\mathrm{RB}_l = \frac{\sum_t \left({n_{l,t}^{est} - n_{l,t}^{true}} \Big/ {n_{l,t}^{true}} \right)}{\tau} \times 100\%$$

Simulation experiments

— Experiment 3: Effect of sample size





Conclusions

- The spatiotemporal model produced unbiased estimates of abundance and fishing mortality spatially
- The spatiotemporal model outperformed a spatially-implicit model when time-varying selectivity caused by spatial heterogeneity in fishing pressure is ignored
- Our modeling approach bridges the gap between species distribution and population dynamic models and provides the opportunity to improve natural resource management and conservation

Discussion

- Adapt to populations with different types of life history through straightforward modifications
- The comparison scenario we show here represents the situation where a strong and persistent gradient of fishing pressure occurs over space and time
- Possible to explicitly model movement
- Selectivity –
    - more biologically interpretable
    - could be corroborated by other field sampling
- The advance comes at the expense of greater data requirements

ACKNOWLEDGEMENTS

- We thank K. Kristensen and the developers of Template Model Builder, without which this analysis would not be feasible
- NOAA "Stock Assessment and Analytic Methods" (SAAM) grant
- Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement NA15OAR4320063
- We thank R. Methot for helpful comments on an earlier draft